



# The Effects of Standard and Enhanced Data Sheets and Brief Video Training on Implementation of Conditional Discrimination Training

Linda A. LeBlanc<sup>1,2</sup>  · Layla A. Sump<sup>1</sup> · Justin B. Leaf<sup>3</sup> · Joseph Cihon<sup>3</sup>

Published online: 4 February 2019

© Association for Behavior Analysis International 2019

## Abstract

Grow and LeBlanc (2013) described practice recommendations for conducting conditional discrimination training with children with autism. One recommendation involved using a specially designed datasheet to provide the preset target stimulus for each trial along with counterbalancing the location of stimuli if a three-item array of comparison stimuli. This study evaluated whether the recommended data sheet would lead to higher procedural integrity of counterbalancing trials compared to a standard data sheet (i.e., targets and arrays are not pre-set). Forty behavior therapists from two provider agencies participated. Participants were randomly assigned to either the standard data sheet condition or the enhanced data sheet condition. Participants watched a short video on Grow and LeBlanc's practice recommendations for a matching task and an orientation to the datasheet for the assigned condition, and then implemented the matching task with a confederate serving in the role of the child with autism. The enhanced data sheet resulted in higher accuracy of implementation on counterbalancing than the standard data sheet, with the largest difference for rotation of the target stimulus across trials and for counterbalancing the placement of the correct comparison stimulus in the array.

**Keywords** Autism · Conditional discrimination · Counterbalancing · Data collection · Discrimination training · Matching · Procedural integrity · Staff training

Children diagnosed with developmental disabilities such as autism require a carefully engineered arrangement of their environment to learn to respond effectively to it (Lovaas, 1977). Thus, many of the initial programs in early intensive behavioral intervention (EIBI) curricula are designed to teach children “learning to learn skills” such as, but not limited to, attending, matching, imitation, and responding to the language of others (Lovaas, 2003; Smith, 2001). All of these skills are taught through a procedure called discrimination training, which involves presenting discriminative stimuli and prompts and differentially reinforcing the target response (Lerman, Valentino & LeBlanc, 2016; Smith, 2001). If

discrimination training procedures for teaching these critical skills are not optimally designed and implemented, several problems may emerge that slow the rate of skill acquisition in EIBI (DiGenarro Reed, Reed, Baez & Maguire, 2011; Grow et al., 2009). For example, simple errors in arranging instructional materials in a matching to sample task can inadvertently establish a positional bias if responses to one position result in a differentially higher rate of reinforcement (e.g., the correct stimulus is often in the right position so that selections in that position encounter a higher rate of reinforcement than other positions) (Galloway, 1967; Kangas & Branch, 2008; Mackay, 1991; Sidman, 1992).

A distinction can be made between two types of discriminations: simple and conditional. Simple discriminations include a discriminative stimulus, a particular response (i.e., the behavior), and a reinforcer, usually in the form of praise and tangible items (i.e., the consequence). A conditional discrimination includes a fourth component called a conditional stimulus (Mackay, 1991). There is often an array of visual stimuli, a corresponding conditional stimulus (i.e., the target or sample) that establishes one of the items in the array as the  $S^+$  and the other(s) as the  $S^-$ , the selection from the array (the

---

✉ Linda A. LeBlanc  
linda@lbehavioral.com

<sup>1</sup> Trumpet Behavioral Health, Dublin, CA, USA

<sup>2</sup> LeBlanc Behavioral Consulting, 1006 Dogwood Drive, Golden, CO 80401, USA

<sup>3</sup> Autism Partnership Foundation and Endicott College, Beverly, MA, USA

matching behavior), and delivery of the reinforcer for a correct response (i.e., the consequence). Examples of programs that establish conditional discriminations are visual–visual matching (both identity and non-identity) and auditory–visual matching (e.g., identifying items based on their name, basic features, functions) among others.

Grow and LeBlanc (2013) described practice recommendations for conducting conditional discrimination training (e.g., skills involving matching to sample) using the specific example of auditory–visual discriminations. However, the recommendations are equally pertinent to visual–visual discriminations such as identity matching. The recommendations are based on synthesis of both the basic and applied literature on stimulus control (Green, 2001; Mackay, 1991; Saunders & Williams, 1998), but some of the specific strategies are more practical in nature. For example, the authors suggest using a specially designed datasheet to increase the accuracy of implementation on the most complex of the recommendations (i.e., carefully arrange the antecedent stimuli and required behaviors: counterbalance the visual and/or auditory stimuli). The authors provide sample data sheets that denote the target stimulus for each trial along with a counterbalanced sequence of item location to facilitate compliance with the recommendations, originally presented in a review article by Green (2001), to (1) rotate the target stimulus across trials, (2) conduct an equal number of trials with each stimulus, (3) counterbalance the placement of comparison stimuli, and (4) counterbalance the placement of the correct comparison in the array across trials. In addition, the data sheet can be used to prompt the data collector to conduct a least-to-most prompting assessment probe that can guide the level of prompting providing in subsequent teaching trials.

The suggestion that this type of enhanced data sheet will increase procedural integrity for these specific recommendations compared to a data sheet that does not provide information to guide counterbalancing has not been empirically tested. This recommendation and others (e.g., minimizing inadvertent instructor cues) make practical sense, but need to be empirically tested, as should the combination of all of the recommendations. The enhanced data sheet might improve, worsen, or have no effect on procedural integrity, and it may be too complicated to use without extensive training. If procedural integrity were worsened, skill acquisition might slow down. If the data sheet was too complicated to use without extensive training, the benefits of the counterbalancing might not be worth the additional time invested in training. The current study evaluated whether the enhanced data sheet would lead to higher procedural integrity on these practice recommendations compared to a standard data sheet (i.e., targets and arrays are not pre-set) after brief training. A group design was used to compare the effects of a brief training and the differing data sheets on procedural integrity for the specific practice recommendations related to trial order and counterbalancing.

## Method

### Participants

Participants were 40 individuals (age 22–34 years) employed as direct intervention therapists at two agencies providing applied behavior analytic services to children diagnosed with autism. See Table 1 for a summary of participant demographics. Three participants had completed their masters' degree, 35 participants had completed their bachelors' degree, 1 participant was an undergraduate student and 2 participant did not report his/her educational background. All participants had been employed for at least 4 months to ensure that initial training was completed. Participants were matched on two relevant variables prior to random assignment: duration of experience and the organization that employed them. Within the 40 participants, 20 participants had 4 months to 1 year of experience (i.e., relatively less experience), while the other 20 participants had at least 2 years of experience (i.e., relatively more experience) working directly with individuals with autism and other disabilities and conducting discrete trial training. Participants were recruited and assigned to condition so that the groups would be equal and matched with respect to the number of individuals with relatively less and relatively more experience in data collection and implementation of discrete trial training procedures with children diagnosed with autism. Participants were recruited via email. Participation did not affect employment or performance evaluation and participants were compensated for their time at the typical company rate for training activities.

### Setting and Materials

Sessions took place in the agencies' administrative offices. The room contained one table and three chairs. Materials consisted of a laptop computer, a video camera for scoring purposes, six black and white matching stimuli (4" × 5") (c.10 × 13 cm) (two circles, two triangles, two squares), one writing utensil, and data sheets. The participant sat across the table from the experimenter during the entire session. The experimenter also served as the confederate learner during the performance test. The webcam was positioned behind the experimenter/confederate and facing the table and the participant. Each data sheet (i.e., independent variable) contained instructions and a scoring system for collecting data on prompts (i.e., independent correct, gestural, model and physical prompt) during assessment probes and teaching trials (see Appendix 1). These data sheets were created for the purposes of the current study rather than adopted from either organization. That is, no participants had prior experience with these exact data sheets.

**Table 1** Participant demographics

Data sheet condition	Gender		Age	Educational level			Duration of employment	
	Female	Male	Mean (range)	BA or pursuing BA	Masters	Not reported	4–12 monts	>2 years
Standard	14	6	26 (23–30)	18	2	0	10	10
Enhanced	19	1	26 (22–34)	18	1	1	10	10

**Standard Data Sheet** Prior to the development of the standard data sheet, Board Certified Behavior Analysts were consulted from both provider agencies to gather information on the components that should be included in a trial by trial data sheet (i.e., continuous data collection) that could be used for general matching tasks. Based on the consultations, the experimenters designed the standard data sheet specifically for the purpose of this experiment. The data sheet contained the following components: a place to write the targets, a place to score the assessment probe for that target, a place to score whether the response emitted by the learner was correct for each teaching trial, and a place to score whether the response emitted by the learner was independent or prompted for each teaching trial (top portion of Appendix 1).

**Enhanced Data Sheet** The enhanced data sheet was designed as recommended by Grow and LeBlanc (2013) (see Appendix 1, bottom portion). For instance, the targets and prompt levels were listed on the data sheet for each probe and teaching trial and the targets were systematically rotated across trials. The user circles the stimulus corresponding to the learner's initial response on each assessment probe and teaching trial and scores the prompt required to produce a correct response. Moreover, the placement of the correct target and comparison stimuli were counterbalanced. Finally, there were an equal number of trials for each stimulus.

## Experimental Design

A group design was used to evaluate the effects of the type of data sheet and a brief training on the use of the data sheet on several dependent variables. Participants were matched with respect to organization and level of experience and randomly assigned to either the standard or the enhanced data sheet condition.

## Measurement and Dependent Variables

All teaching sessions were scored from videotape for both primary and secondary scoring. The observer used a blank modified version of the enhanced data sheet to record the information observed in each teaching trial. The observer scored the teaching trial for the target stimulus (i.e., circle, square, triangle) along with the positional placement of the

three comparison stimuli and which was the discriminative stimulus (e.g., circle-left, triangle-center, square-right; target was triangle-center).

The experimenters used the scored teaching trials from the session to evaluate correspondence of the teaching procedures with the recommended practices for counterbalancing in Grow and LeBlanc (2013) and Green (2001). The four dependent variables matched the recommendations made by Grow and LeBlanc (2013) in designing the data sheet: (1) rotate the target stimulus across trials, (2) conduct an equal number of trials with each target stimulus, (3) counterbalance the placement of comparison stimuli, and (4) counterbalance the placement of the correct comparison in the array across trials. The experimenters coded: (1) the rotation of target across teaching trials (i.e., was each target taught in each three-trial block or was there duplication of the target stimulus within a 3-trial block), (2) the number of teaching trials per target, (3) the percentage of trials that each comparison stimulus appeared in each position (i.e., left, middle, right) of the array, and (4) the percentage of trials that the target stimulus appeared in each position of the array.

The first dependent variable was the percentage accuracy in rotating the target stimulus across teaching trials (i.e., one trial with each target in each three-trial block of teaching trials). The second and third of each three-trial block was scored for whether the target stimulus was a duplicate of either of the prior trials. The number of trials with no duplication was divided by the total number of trials scored (i.e.,  $n = 6$ ).

The second dependent variable was accuracy in conducting an equal number of teaching trials with each target stimulus (i.e., three trials of each of the three targets in the 9-trial session). This was calculated as the absolute value of three minus the number of times the stimulus was the target for each of the three stimuli {i.e.,  $[ABS(3 - \# \text{ square trials})] + [ABS(3 - \# \text{ circle trials})] + [ABS(3 - \# \text{ triangle trials})]$ }. Perfect implementation generated a total of 0 and all trials with the same stimulus generated a score of 12. The average score for the group was divided by 12 to generate the error percentage and the reciprocal is the accuracy percentage.

The third dependent variable was the percentage of accuracy in counterbalancing the placement of the comparison stimuli in the array. In a fully counterbalanced arrangement of a 3-stimulus array, each stimulus appears in each of the three positions exactly 33% of the time. This error percentage was

calculated as the absolute value of 33 minus the percentage of trials that the stimulus appeared in each position for each of the three stimuli {i.e.,  $[ABS(33 - \% \text{ square stimulus left position})] + [ABS(33 - \% \text{ square center position})] + [ABS(33 - \% \text{ square right position})] + [ABS(33 - \% \text{ circle stimulus left position})] + [ABS(33 - \% \text{ circle center position})] + [ABS(33 - \% \text{ circle right position})] + [ABS(33 - \% \text{ triangle stimulus left position})] + [ABS(33 - \% \text{ triangle center position})] + [ABS(33 - \% \text{ triangle right position})]$ }. That is, fully correct implementation generated an error score of 0 and errors generated scores up to 297. The accuracy percentage for the group was calculated as the reciprocal of the average error score divided by 3.

The fourth dependent variable was the percentage of accuracy in counterbalancing the placement of the correct comparison stimulus according to the recommendations of Grow and LeBlanc (2013). In a fully counterbalanced arrangement of a 3-stimulus array, each stimulus serves as the discriminative stimulus one time in each of the three positions across the nine teaching trials. This error percentage was calculated for teaching trials as the absolute value of one minus the number of times that the stimulus was the target for each position for each of the three stimuli {i.e.,  $[ABS(1 - \# \text{square discriminative stimulus left position})] + [ABS(1 - \# \text{square discriminative stimulus center position})] + [ABS(1 - \# \text{square discriminative stimulus right position})] + [ABS(1 - \# \text{circle discriminative stimulus left position})] + [ABS(1 - \# \text{circle discriminative stimulus center position})] + [ABS(1 - \# \text{circle discriminative stimulus right position})] + [ABS(1 - \# \text{triangle discriminative stimulus left position})] + [ABS(1 - \# \text{triangle discriminative stimulus center position})] + [ABS(1 - \# \text{triangle discriminative stimulus right position})]$ }. Thus, perfect implementation resulted in an error score of zero and the highest possible score was 12. The average error score for the group was divided by 12 and the reciprocal served as the accuracy percentage.

### Interobserver Agreement and Procedural Integrity

A second independent observer scored 30% of the videos across the two conditions. An agreement was defined as both observers recording exactly the same information for a given component of a trial (i.e., both observers recorded that circle was placed in the left position of an array). A disagreement was defined as any discrepancy in the scoring of any component of trial (e.g., one recorded that the target stimulus was square while the other recorded circle). Mean agreement was 99.9% (range, 98.5–100%) and it was calculated by dividing the number of agreements by the total agreements plus disagreements and multiplying by 100%.

In addition, an observer scored video of 20% of sessions across both conditions to evaluate the accuracy of procedural implementation. The observer scored the behavior of the experimenter/confederate against a checklist that specified behaviors that needed to occur one time per session (i.e., was the

correct video for the condition played, were all materials placed correctly) or on each trial (i.e., did the experimenter/confederate respond to trials according to the script, did the experimenter/confederate refrain from reinforcing or correcting the participant's teaching behavior, did the experimenter/confederate refrain from answering any questions that were posed) to evaluate the accuracy of the experimenter/confederate's procedural implementation. Procedural integrity was 100%.

### Procedure

Prior to beginning the session, the participants filled out a background form to gather participants' information including age, gender, job title, educational background, and duration of experience working with individuals with disabilities. Participants then viewed a video with detailed instruction on (1) the Grow and LeBlanc (2013) recommendations for conducting a matching task (identical for both experimental conditions), and (2) an orientation to the data sheet for the assigned experimental condition. The video was viewed only once with no opportunity for additional review. The videos for the two conditions were identical on the instructions for the teaching procedure and differed only in the instructions for the use of the data sheet.

The video was narrated by the third author and contained graphic illustration along with the instructions. The portion of the video on the matching task instructed participants to place at least three stimuli in a comparison array for each trial, to arrange the stimuli out of the view and reach of the learner with the stimuli evenly spaced, and to present the array of stimuli simultaneously rather than one at a time. The video also instructed participants to vary targets across trials, to ensure that the comparison stimuli appeared in each array position an equal number of times, and to ensure that the position of the correct stimulus was equal for each position (i.e., correct on the right the same number of times as correct in the middle or left). Finally, the video instructed participants to start the session by conducting a single least-to-most probe for each target to identify the controlling prompt that would be used in subsequent most-to-least prompting teaching trials.

Following the instructions for the matching task, the video described the data sheet for the assigned experimental condition. The video for the standard data sheet condition instructed participants to write the name of each target that was going to be taught in the session in one of the boxes. Then, participants were told to conduct the probe and teaching trials next, by arranging the stimuli with even spacing and out of the confederate's view, and by presenting the target stimulus while sliding all pictures across the table at the same time. The participants were instructed to score the response of the confederate for each target by indicating whether the response was correct (+), incorrect or no response (–), as well as the prompt level used for each trial. Moreover, participants were told to use

least-to-most prompting during the probe trials to identify the type of prompt to use in the subsequent teaching trials. The participants were instructed to use most-to-least prompting for the teaching trials, beginning with the identified prompt from the probe trial. Participants were also told that the responses in the teaching trials would almost always be correct because of the use of most to least teaching. Next, participants were asked to conduct a total of nine teaching trials across the three targets during the session (i.e., the post-test). They were informed that not every teaching trial slot in the data sheet would be filled out (total video duration = 8 min).

The video for the enhanced data sheet condition instructed participants to arrange the stimuli out of the confederate's view, and to present the target sample stimulus while sliding the three pictures across the table. Participants were told to use least-to-most prompting during the probe trials to identify the type of prompt to use in the subsequent teaching trials. They were also told to circle the stimulus that the learner first pointed to, even if that response was wrong and to circle the prompt level that was eventually required to produce the correct response. The participants were instructed to use most-to-least prompting for the subsequent teaching trials, beginning with the identified prompt from the probe trial. Then, the participants were told to circle the stimulus that the learner first pointed to, even if that response was wrong and to circle the prompt level that was needed to evoke correct responding. Finally, participants were also told that the responses in the teaching trials would almost always be correct because of the use of most to least teaching (total video duration = 9 min). Each video ended with a brief introduction to the performance post-test (e.g., teach an adult confederate to perform a matching task using the materials on the table and do your best).

The performance post-test started when the participant reached for the matching task materials and ended 1 s after the end of the last trial with the confederate. A confederate learner followed a predetermined script to ensure that all participants had the opportunity to respond to the same number of learner errors and correct responses on the part of the learner. The script for the response for each trial was displayed on a laptop computer that faced the confederate out of the view of the participant. The confederate did not respond to any questions that the participants may have asked, and no praise or reinforcement was delivered. That is, this was the performance post-test rather than a rehearsal and feedback opportunity.

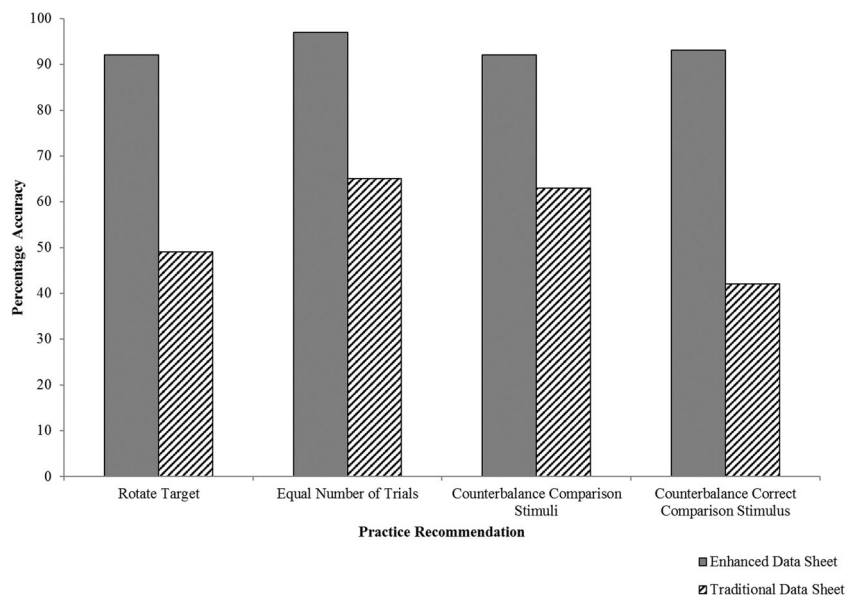
## Results

The primary analysis was a comparison of the two data sheet (and training on the data sheet) conditions (i.e., standard vs. enhanced). See Fig. 1 for the group mean percentage accuracy for each of the practice recommendations for the enhanced

and standard data sheet (plus training) conditions. Participants in the enhanced data sheet condition rotated the targets across trials with 92% accuracy (range, 16–100%), while participants in the standard data sheet group had a mean score of 49% (range, 0–100%). A *t* test confirmed the difference in accuracy of target rotation was significant,  $t(38) = 4.24, p < 0.00$ . When it comes to conducting equal number of trials, the enhanced data sheet group mean accuracy was 97% (range, 83–100%) and the standard data sheet group was 69% (range, 0–100%). A *t* test confirmed the difference in number of trials was significant,  $t(38) = 3.88, p < 0.00$ . Participants in the enhanced data sheet (plus training) condition counterbalanced the comparison stimuli with 92% accuracy (range, 60–100%), and participants in the standard data sheet (plus training) condition had the mean score was 63% (range, 48–85%). A *t* test confirmed the difference in accuracy of counterbalancing the comparison stimuli was significant  $t(38) = 6.06, p < 5.17$ . Finally, participants in the enhanced data sheet condition counterbalanced the correct comparison stimulus with 93% accuracy (range, 63–100%) versus 42% accuracy (range, 25–85%) for participants in the standard data sheet condition. A *t* test confirmed the difference in accuracy of counterbalancing the placement of the correct comparison stimulus was significant,  $t(38) = 7.68, p < .00$ . We also examined combined accuracy averaged across the four recommendations. Participants who used the enhanced data sheet implemented the four practice recommendations with higher accuracy (i.e., 93.5% accuracy across all variables) than participants using the standard data sheet (i.e., 57.5% mean accuracy). No *t* test was conducted for this metric as it is a combination of the metrics listed above.

The percentage of participants who implemented each of the four practice recommendations at 100% accuracy is depicted on Table 2. This is a more stringent success criterion than depicted in Fig. 1. Overall, more participants in the enhanced data sheet group (68.5%) implemented the practice recommendations perfectly when compared to participants in the standard data sheet group (18%). The first practice recommendation evaluated was rotating the target stimulus across trials, where 82% of the participants in the enhanced data sheet group performed this recommendation perfectly compared to 27% of the participants in the standard data sheet group. The second practice recommendation evaluated was conducting an equal number of trials with each target stimulus in the set, where 64% of the participants in the enhanced data sheet group performed this recommendation perfectly compared to 45% of the participants in the standard data sheet group. The third and fourth practice recommendations evaluated were counterbalancing the placement of the comparison stimuli in the array and counterbalancing the placement of the correct comparison stimulus, and 64% of the participants in the enhanced data sheet group performed this recommendation perfectly compared to 0% of the participants in the standard data sheet group across both recommendations.

**Fig. 1** The group mean percentage accuracy for each of the practice recommendations is presented for the enhanced data sheet condition (*solid bar*) and the standard data sheet condition (*hashed bar*)



Follow-up analyses were conducted to determine if reported years of experience impacted the results. The results were all well above 80% accuracy and were identical or nearly identical for the two subgroups of participants in the enhanced data sheet condition (e.g., 94/91, 96/98). For the standard data sheet condition, the mean scores of the two sub-groups varied a little more (i.e., 3–18 percentage point spread) and all were well below 80% (see Table 3). The largest discrepancy was for an equal number of teaching trials with each target due to a few of the more experienced participants conducting all trials with a single target. The matching procedure for level of experience ensured that each condition had an equal number of more experienced and less experienced participants, minimizing the chances that assignment bias of more experienced participants to one group accounted for the observed differences in the accuracy of implantation for the two groups.

## Discussion

Grow and LeBlanc (2013) described practice recommendations for conducting conditional discrimination training with children with autism that might minimize the likelihood of faulty stimulus control (Sidman, 1992) that sometimes occurs with simple to conditional discrimination training, particularly

for the earliest learners. Many of the recommendations were based on the experimental literature on stimulus control (Green, 2001), and each recommendation has subsequently been demonstrated and replicated either in isolation or in some combination in applied studies with children with autism (Farber, Dickson, & Dube, 2017; Fisher, Retzlaff, Akers, DeSouza & Kaminski, 2019; Grow, Carr, Kodak, Jostad & Kisamore, 2011; Grow, Kodak & Carr, 2014; Vedora & Grandeliski, 2015). One recommendation in particular involved using a specially designed datasheet to provide the preset target stimulus and a counterbalanced three-item array of comparison stimuli for each trial. Grow and LeBlanc suggested that a well-designed data sheet could assist instructors with implementation of counterbalancing and provided a sample datasheet with pre-set targets and arrays, but the recommendation did not have any empirical support to demonstrate such an effect. The current study evaluated whether the recommended datasheet, in addition to brief training in the practice recommendations and use of the datasheet, would result in higher procedural integrity on the counterbalancing recommendations than a datasheet without preset arrays and similar brief instruction.

Implementation of each practice recommendation was below 80% accuracy with the standard data sheet even when the performance opportunity occurred immediately after training.

**Table 2** The percentage of participants in each condition with perfect implementation for each of the for recommendations

Practice recommendation	Standard data sheet (%)	Enhanced data sheet (%)
Rotate target stimulus across trials	27	82
Equal number of trials with each target stimulus	45	64
Counterbalance placement of comparison stimuli	0	64
Counterbalance placement of correct comparison stimulus	0	64

**Table 3** The mean percentage accuracy for the standard data sheet condition for all participants, less experienced participants, and more experienced participants

Practice recommendation	All participants	Less experienced	More experienced
Rotate target stimulus across trials	49	51	46
Equal number of trials with each target stimulus	65	77	59
Counterbalance placement of comparison stimuli	63	62	65
Counterbalance placement of correct comparison stimulus	42	39	46

The enhanced data sheet resulted in participants having a higher accuracy of implementation (i.e., above 90%) for each of the recommended practices than the participants using the standard data sheet. The largest discrepancies across groups were in the tasks of rotating the target stimulus across trials and for counterbalancing the placement of the correct comparison stimulus in the array. Surprisingly, participants were able to do moderately well on counterbalancing the stimulus array (i.e., each stimulus appearing in each position on 1/3 of trials) even without the enhanced data sheet. This might seem like one of the more difficult recommendations to implement without a structured aid; however, participants developed strategies such as moving the stimuli one position to the right or left for each subsequent trial and achieved over 60% accuracy with this strategy.

In addition, the results suggest that the enhanced data sheet (plus the training) eliminated the differences in performance for more and less experienced participants. For the standard data sheet (plus training), differences remained between more and less experienced participants. There may be very good reasons for the errors observed in the standard data sheet and the resulting instructional programming might well still be effective with children with autism. For example, more experienced therapists may have elected to focus on a single target based on an error response by the confederate. These participants may very well have been following a reasonable and systematic instructional strategy that simply did not correspond with the training that was provided.

These findings provide support for the recommendation that training and an enhanced data sheet increases procedural integrity for these specific practice recommendations, should one choose to implement the recommendations as stated by Grow and LeBlanc (2013). In addition, only a brief training was required for each type of data sheet. This finding supports other research findings that use of this set of recommendations and this data sheet can be taught relatively quickly either through live training or e-learning (Geiger, LeBlanc, Hubik, Jenkins & Carr, 2018). Practice recommendations are often based on a combination of synthesized empirical evidence and practical clinical experience (e.g., Grow & LeBlanc, 2013; Lamarca & LaMarca, 2018; Leaf, Cihon, Leaf, McEachin & Taubman, 2016). It is important to experimentally evaluate those that do not yet have experimental support, as was the case with the recommendation to use a structured

data sheet. It should also be noted that the demonstration of higher procedural integrity is only of value if the correct implementation of the procedure actually mitigates the incorrect learning strategies that it is intended to prevent. Since confederates were used in this training study as well as in Geiger et al., 2018, we cannot draw any conclusions about the impact of this type of instruction on the skill acquisition progress of children with autism. Additional studies could examine the effects of this type of instruction, or differential accuracy in implementation on skill acquisition of children with autism.

Several cautions are warranted in interpreting these data. First, this study only evaluated the effects of the use of two types of data sheets and brief training on compliance with a set of practice recommendations. This study should not be taken as evidence that the practice recommendations themselves are valid. Others have provided a different set of practice recommendations advocating for a more flexible and progressive approach to discrimination training (e.g., Leaf et al., 2016). As with most sets of practice recommendations, the term “best” practice should be avoided as it is rarely the case that the recommendations have been empirically evaluated in comparison to every other possible set of recommendations. It is also the case that new research may identify conditions under which the original practice recommendations are not optimal. For example, recent studies have in addition examined one of the recommendations not included in this study (i.e., present an auditory stimulus before the visual stimuli and continue to repeat the auditory stimulus every 2 s until selection). For some children with autism, it is more advantageous to present the auditory stimulus first (Petursdottir & Aguilar, 2016), while for others it is better to present the comparison stimulus first (Schneider, Devine, Aguilar & Petursdottir, 2018). Second, the current evaluation was conducted in a role play context with a confederate rather than with children with autism. Although it would have been optimal to have conducted every post-test with a child with autism, it was not practical to identify multiple children across sites who all needed exactly the skill targets included in this study. Finally, the current evaluation used a two-group post-test group design. Thus, no pre-test comparison was conducted. Future studies might use a two-group pre-test, post-test design.

In this study, the data sheet differed across conditions and we also trained the participants in the use of the data sheet and the desired implementation strategy. The training videos were

identical with respect to the training in the procedure and differed only in the description of the data sheet and how to use it. The fact that the groups experienced the exact same training in the procedures (i.e., this variable did not vary across the groups) suggests that the obtained differences are due to the data sheet used in each condition. The training was included in order to follow recommended practices for human services that indicate that training should be provided for staff prior to expecting high accuracy of performance (Parsons, Rollyson & Reid, 2012). However, the inclusion of the brief training does not allow us to identify if different data sheets would produce differential accuracy if we did not provide training in how to do the procedure or use the data sheet. Without training as a translational study (i.e., in applied settings) you should always provide training).

While these findings suggest that a specially designed data sheet and a brief instructional video can enhance procedural integrity, other studies are needed. The current study should be replicated with both paper and pencil data and with electronic data collection tools. Electronic data collection tools can have the advantage of using algorithms to adjust to error patterns, so that, if a child makes an error to a specific stimulus or position on a preceding trial, the position assignments for the

subsequent trial are immediately altered to ensure that reinforcement is not available for that same stimulus or position. Future research studies could compare various combinations of the Grow and LeBlanc (2013) recommendations to determine which are necessary and under which conditions. Studies might also compare various models of programming, such as the recently published discrimination training curriculum from New England Center for Children that is based on the work of Green (2001) and many others (MacDonald & Langer, 2018), and the recently published recommendations for flexible programming (Leaf et al., 2016). In addition, these studies could be conducted in the context of programming with children with autism instead of with confederates.

**Author Note** The authors thank Breanna Mottau for assistance with data collection and analysis.

**Compliance with Ethical Standards**

The authors of this manuscript declare no conflict of interest regarding this manuscript. No funding was associated with the current study.

Ethical Approval: All procedures were performed in accordance with the ethical standards of the institutional review committee and with the 1964 Helsinki declaration and its later amendments.

**Appendix 1. Enhanced (top) and Standard (bottom) data sheets**

Participant ID: \_\_\_\_\_ Session date: \_\_\_\_\_

**Instructions:**

1. **Arrange** the stimuli as indicated per trial
2. Present the sample (**bolded one is the target** for each trial).
3. **Circle** the stimulus that the learner **FIRST** points to – even if it is wrong.
4. **Circle** the **prompt level** required to get the correct response.
5. The first 3 trials (highlighted in gray) are **LEAST TO MOST PROBES**.
6. The rest are **MOST TO LEAST TEACHING TRIALS**.

**Prompt Scoring Key:**

I = Independent Correct Response    G = Gestural Prompt  
 PP = Partial Physical Prompt        FP = Full Physical Prompt

		<b>Left</b>	<b>Center</b>	<b>Right</b>	<b>Prompt Req.</b>
<b>Probes</b>	1	<b>Triangle</b>	Circle	Square	I G PP FP
	2	Circle	<b>Square</b>	Triangle	I G PP FP
	3	Square	Triangle	<b>Circle</b>	I G PP FP
<b>Teaching Trials</b>	4	Circle	<b>Square</b>	Triangle	FP PP G I
	5	Square	Triangle	<b>Circle</b>	FP PP G I
	6	<b>Triangle</b>	Circle	Square	FP PP G I
	7	<b>Circle</b>	Square	Triangle	FP PP G I
	8	Square	<b>Triangle</b>	Circle	FP PP G I
	9	Triangle	Circle	<b>Square</b>	FP PP G I
	10	Circle	Square	<b>Triangle</b>	FP PP G I
	11	<b>Square</b>	Triangle	Circle	FP PP G I
	12	Triangle	<b>Circle</b>	Square	FP PP G I



Participant ID: \_\_\_\_\_ Session date: \_\_\_\_\_

**Instructions:**

1. Write the targets that will be taught this session in the target column.
2. Score the response of the student in the probe or teaching trial column:
  - a. Write + for an independent correct response
  - b. Write – for an incorrect response
3. Write the prompt level required to get the correct response in the “Prompt used” row.
4. The first trial per target (highlighted in gray) is a **LEAST TO MOST PROBE**.
5. The rest are **MOST TO LEAST TEACHING TRIALS**.

**Prompt Scoring Key:**

I = Independent Correct Response    G = Gestural Prompt  
 PP = Partial Physical Prompt        FP = Full Physical Prompt

Target		P	1	2	3	4	5	6	7	8	9
	+ or -										
	Prompt used										
	+ or -										
	Prompt used										
	+ or -										
	Prompt used										

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- DiGennaro Reed, F. D., Reed, D. D., Baez, C. N., & Maguire, H. (2011). A parametric analysis of errors of commission during discrete-trial training. *Journal of Applied Behavior Analysis*, *44*, 611–615. <https://doi.org/10.1901/jaba.2011.44-611>.
- Farber, R. S., Dickson, C. A., & Dube, W. V. (2017). Reducing overselective stimulus control with differential observing responses. *Journal of Applied Behavior Analysis*, *50*, 87–105. <https://doi.org/10.1002/jaba.363>.
- Fisher, W. W., Retlaff, B. J., Akers, J. S., DeSouza, A. A., & Kaminski, A. J. (2019). Establishing initial auditory-visual conditional discriminations and emergence of initial tacts in young children with autism spectrum disorder. *Journal of Applied Behavior Analysis*.
- Galloway, C. (1967). Modification of a response bias through a differential amount of reinforcement. *Journal of the Experimental Analysis of Behavior*, *10*, 375–382. <https://doi.org/10.1901/jeab.1967.10-375>.
- Geiger, K. B., LeBlanc, L. A., Hubik, K., Jenkins, S. R., & Carr, J. E. (2018). Live training versus e-learning to teach implementation of listener response programs. *Journal of Applied Behavior Analysis*, *51*, 220–235. <https://doi.org/10.1002/jaba.444>.
- Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, *16*, 72–85. <https://doi.org/10.1177/108835760101600203>.
- Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice*, *6*, 56–75. <https://doi.org/10.1007/BF03391791>.
- Grow, L. L., Carr, J. E., Gunby, K. V., Charania, S. M., Gonsalves, L., Ktaech, I. A., & Kisamore, A. N. (2009). Deviations from prescribed prompting procedures: Implications for treatment integrity. *Journal of Behavioral Education*, *18*, 142–156. <https://doi.org/10.1007/s10864-009-9085-6>.
- Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, *44*, 475–498. <https://doi.org/10.1901/jaba.2011.44-475>.
- Grow, L. L., Kodak, T. M., & Carr, J. E. (2014). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders: A systematic replication. *Journal of Applied Behavior Analysis*, *47*, 600–605. <https://doi.org/10.1002/jaba.141>.
- Kangas, B. D., & Branch, M. N. (2008). Empirical validation of a procedure to correct position and stimulus biases in matching to sample. *Journal of the Experimental Analysis of Behavior*, *90*, 103–112. <https://doi.org/10.1901/jeab.2008.90-103>.
- LaMarca, V., & LaMarca, J. (2018). Designing receptive language programs: Pushing the boundaries of research and practice. *Behavior Analysis in Practice*. <https://doi.org/10.1007/s40617-018-0208-1>.
- Leaf, J. B., Cihon, J. H., Leaf, R., McEachin, J., & Taubman, M. (2016). A progressive approach to discrete trial teaching: Some current guidelines. *International Electronic Journal of Elementary Education*, *9*, 361–372.
- Lerman, D. L., Valentino, A., & LeBlanc, L. A. (2016). Discrete trial training. In R. Lang, T. Hancock, & N. N. Singh (Eds.), *Early intervention for children with autism spectrum disorders* (pp. 47–84). New York: Springer.

- Lovaas, O. I. (1977). *The autistic child: Language development through behavior modification*. New York: Irvington.
- Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques*. Austin: PRO-ED.
- MacDonald, R., & Langer, S. (2018). *Teaching essential discrimination skills to children with autism: A practical guide for parents and educators*. Bethesda: Woodbine House.
- Mackay, H. A. (1991). Conditional stimulus control. In I. H. Iversen & K. A. Lattal (Eds.), *Techniques in the behavioral and neural sciences: Vol. 6. Experimental analysis of behavior* (Part 1 (pp. 301–350). Amsterdam: Elsevier.
- Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice*, 5(2), 2–11.
- Petursdottir, A. I., & Aguilar, G. (2016). Order of stimulus presentation influences children's acquisition in receptive identification tasks. *Journal of Applied Behavior Analysis*, 49, 58–68. <https://doi.org/10.1002/jaba.264>.
- Saunders, K. J., & Williams, D. C. (1998). Stimulus-Control Procedures. In K. A. Lattal & M. Perrone (Eds.), *Handbook of research methods in human operant behavior*. Boston: Springer.
- Schneider, K. A., Devine, B., Aguilar, G., & Petursdottir, A. I. (2018). Stimulus presentation order in receptive identification tasks: A systematic replication. *Journal of Applied Behavior Analysis*, 51, 634–646. <https://doi.org/10.1002/jaba.459>.
- Sidman, M. (1992). Adventitious control by the location of the comparison stimuli in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 58, 173–182. <https://doi.org/10.1901/jeab.1992.58-173>.
- Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and other Developmental Disabilities*, 16, 86–92. <https://doi.org/10.1177/108835760101600204>.
- Vedora, J., & Grandeliski, K. (2015). A comparison of methods for teaching receptive language to toddlers with autism. *Journal of Applied Behavior Analysis*, 48, 188–193. <https://doi.org/10.1002/jaba.167>.